CLAIMS

1

1. A method of incrementally updating precision and recall curves in a k nearest
neighbor database, said database including original documents, categories, category
assignments for the original documents, and category scores for the original
documents, the method including:

retaining for the original documents a list of their m nearest neighbors and
corresponding similarity scores, wherein m>k;

adding or deleting one or more original documents and their category
assignments;

identifying the documents influenced by the adding or deleting;

updating one or more category scores of the influenced documents; and

computing precision and recall curves for the categories having updated category
scores.

2. A method of incrementally updating precision and recall curves in a k nearest
neighbor database, said database including original documents, categories, category
assignments for the original documents, and category scores for the original
documents, the method including:

retaining for the original documents a list of their m nearest neighbors and
corresponding similarity scores, wherein m>k;

adding or deleting one or more category assignments to one or more original
documents;

updating category scores of the documents influenced by the adding or deleting
of one or more category assignments, for at least the categories to which the
category assignments were added or deleted; and

computing precision and recall curves for the categories having updated category
scores.

1   3. A method of incrementally adding category assignments to particular original

2   documents in a k nearest neighbor database, said database including original

3   documents, categories, category assignments for the original documents, and category

4   scores for the original documents, the method including:

5   retaining for the original documents a first list of their k nearest neighbors and

6   corresponding similarity scores;

7   retaining for the original documents a second list of m-k additional nearest

8   neighbors and corresponding similarity scores;

9   adding one or more category assignments for one or more particular original

10   documents;

11   computing category scores for the particular original documents and a

12   predetermined number of nearest neighbors of the particular original documents,

13   for those categories to which the category assignments are added, based on the

14   retained similarity scores; and

15   computing precision and recall curves for the categories to which the category

16   assignments are added.

1   4. A method of incrementally adding one or more documents to a k nearest

2   neighbor database, said database including original documents, categories, category

3   assignments for the original documents, and category scores for the original

4   documents, the method including:

5   retaining for the original documents a first list of their k nearest neighbors and

6   corresponding similarity scores;

7   retaining for the original documents a second list of m-k additional nearest

8   neighbors and corresponding similarity scores;

9   adding one or more documents;

10   calculating similarity scores between the added documents, and the added and

11   original documents;

12   modifying the retained first and second nearest neighbor lists for a predetermined

13   number of nearest neighbors of the added documents;

14   adding category assignments for the added documents;

15   computing one or more category scores for the added documents and the

16   predetermined number of nearest neighbors of the added documents, based on the

17   retained and calculated similarity scores; and

18   computing precision and recall curves for the categories to which the category

19   assignments are added.

1   5. A method of incrementally deleting category assignments from particular

2   documents in a k nearest neighbor database, said database including original

3   documents, categories, category assignments for the original documents, and category

4   scores for the original documents, the method including:

5   retaining for the original documents a first list of their k nearest neighbors and

6   corresponding similarity scores;

7   retaining for the original documents a second list of m-k additional nearest

8   neighbors and corresponding similarity scores;

9   deleting one or more of the category assignments for one or more particular

10   original documents;

11   computing category scores for the particular original documents and a

12   predetermined number of nearest neighbors of the particular original documents,

13   for those categories from which the category assignments are deleted, based on

14   the retained similarity scores; and

15   computing precision and recall curves for the categories from which the category

16   assignments are deleted.

1   6. A method of incrementally deleting documents from a k nearest neighbor

2   database, said database including original documents, categories, category

3   assignments for the original documents, and category scores for the original

4   documents, the method including:

5   retaining for the original documents a first list of their k nearest neighbors and

6   corresponding similarity scores;

7   retaining for the original documents a second list of m-k additional nearest

8   neighbors and corresponding similarity scores;

9   deleting one or more of the original documents and corresponding category

10   assignments from the database;

11   deleting the deleted documents from the retained first and second nearest

12   neighbor lists for a predetermined number of nearest neighbors of the deleted

13   documents;

14   computing one or more category scores for a predetermined number of nearest

15   neighbors of the deleted documents, based on the retained similarity scores; and

16   computing precision and recall curves for the categories in which the deleted

17   documents had category assignments.

1   7. A method of incrementally adding category assignments to particular original

2 documents in a k nearest neighbor database, said database including original

3 documents, categories, category assignments for the original documents, and category

4 scores for the original documents, the method including:

5   retaining for the original documents a first list of their k nearest neighbors and

6   corresponding similarity scores;

7   creating an influence list of original documents having a particular original

8   document among their k nearest neighbors;

9   adding one or more category assignments for one or more particular original

10   documents;

11   identifying influenced original documents from the influence list for the

12   particular original documents to which the category assignments are added;

13   computing category scores of the influenced original documents and of the

14   particular original documents, for those categories to which the category

15   assignments are added, based on the retained similarity scores; and

16   computing precision and recall curves for the categories to which the category

17   assignments are added.

1   8. A method of incrementally adding one or more documents to a k nearest

2 neighbor database, said database including original documents, categories, category

3 assignments for the original documents, and category scores for the original

4 documents, the method including:

5   retaining for the original documents a first list of their k nearest neighbors and

6   corresponding similarity scores;

7   creating an influence list of those original documents having certain original

8   documents among their k nearest neighbors;

9   adding one or more documents to the database;

10   calculating similarity scores between the added documents, and the added and

11   original documents;

12   updating the retained first list of k nearest neighbors to include the added

13   documents;

14   updating the influence list to include the added documents;

15   adding category assignments for the added documents;

16   computing one or more category scores of the added and original documents

17   influenced by the category assignments, based on the retained and calculated

18   similarity scores; and

19   computing precision and recall curves for the categories to which the category

20   assignments are added.

1   9. A method of incrementally deleting category assignments from particular

2 documents in a k nearest neighbor database, said database including original

3   documents, categories, category assignments for the original documents, and category

4   scores for the original documents, the method including:

5   retaining for the original documents a first list of their k nearest neighbors and

6   corresponding similarity scores;

7   creating an influence list of those original documents having certain original

8   documents among their k nearest neighbors;

9   deleting one or more category assignments for one or more particular original

10   documents;

11   identifying influenced original documents from the influence list for the

12   particular original documents from which the category assignments are deleted;

13   computing category scores of the influenced original documents and of the

14   particular original documents for those categories from which the category

15   assignments are deleted, based on the retained similarity scores; and

16   computing precision and recall curves for the categories from which the category

17   assignments are deleted.

1   10. A method of incrementally deleting one or more documents to a k nearest

2   neighbor database, said database including original documents, categories, category

3   assignments for the original documents, and category scores for the original

4   documents, the method including:

5   retaining for the original documents a first list of their k nearest neighbors and

6   corresponding similarity scores;

7   retaining for the original documents a second list of m-k additional nearest

8   neighbors and corresponding similarity scores;

9   creating an influence list of those original documents having certain original

10   documents among their k nearest neighbors;

11   deleting one or more documents from the database and corresponding category

12   assignments;

13   updating the retained first and second lists of m nearest neighbors to delete the

14   deleted documents;

15   updating the influence list to delete the deleted documents;

16   computing one or more category scores of the original documents influenced by

17   the deleted documents, based on the retained similarity scores; and

18   computing precision and recall curves for the categories in which the deleted

19   documents had category assignments.